

A NOVEL GROWTH FACTOR AND A GENETIC SEQUENCE  
ENCODING SAME

5 The present invention relates generally to an isolated molecule having vascular endothelial growth factor-like properties and to a genetic sequence encoding same. The molecule will be useful in the development of a range of therapeutics and diagnostics useful in the treatment, prophylaxis and/or diagnosis of conditions requiring enhanced or diminished vasculature and/or vascular permeability. The molecule of the present  
10 invention is also a useful effector of primary and central neurons and is capable of inducing astroglial proliferation.

Bibliographic details of the publications referred to by author in this specification are collected at the end of the description. Sequence Identity Numbers (SEQ ID NOs.) for  
15 the nucleotide and amino acid sequences referred to in the specification are defined following the bibliography.

Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to  
20 imply the inclusion of a stated element or integer or group of elements or integers but not the exclusion of any other element or integer or group of elements or integers.

Vascular endothelial growth factor (hereinafter referred to as "VEGF"), also known as vasoactive permeability factor, is a secreted, covalently linked homodimeric glycoprotein  
25 that specifically activates endothelial tissues (Senger *et al.*, 1993). A range of functions have been attributed to VEGF such as its involvement in normal angiogenesis including formation of the corpus luteum (Yan *et al.*, 1993) and placental development (Sharkey *et al.*, 1993), regulation of vascular permeability (Senger *et al.*, 1993), inflammatory angiogenesis (Sunderkotter *et al.*, 1994) and autotransplantation (Dissen *et al.*, 1994) and  
30 human diseases such as tumour promoting angiogenesis (Folkman & Shing, 1992), rheumatoid arthritis (Koch *et al.*, 1994) and diabetes related retinopathy (Folkman & Shing, 1992).

VEGF is, therefore, an important molecule making it a potentially valuable target for research into therapeutics, prophylactics and diagnostic agents based on VEGF or its activities. There is also a need to identify homologues or otherwise related molecules for use as an alternative to VEGF or in conjunction with VEGF.

5

In work leading up to the present invention, the inventors sought the multiple endocrine neoplasia type I susceptibility gene (MEN1). Surprisingly, the inventors discovered that a genetic sequence excluded as a candidate for the MEN1 gene was nevertheless a new growth factor having some similarity to VEGF. Furthermore, the growth factor of the  
10 present invention is an effector molecule for primary and central neurons.

Accordingly, one aspect of the present invention comprises a biologically isolated proteinaceous molecule comprising a sequence of amino acids which:

- (i) is at least about 15% similar to the amino acid sequence set forth in SEQ ID  
15 NO:2; and
- (ii) is at least 5% dissimilar to the amino acid sequence set forth in SEQ ID NO:2.

Another aspect of the present invention provides a biologically isolated proteinaceous molecule having the following characteristics:

- 20 (i) comprises an amino acid sequence having at least about 15% similarity but at least about 5% dissimilarity to all or part of the amino acid sequence set forth in SEQ ID NO:2;
- (ii) exhibits at least one property in common with VEGF.

25 A related aspect of the present invention contemplates a biologically isolated proteinaceous molecule having the following characteristics:

- (i) comprises an amino acid sequence having at least about 15% similarity but at least about 5% dissimilarity to the amino acid sequence set forth in SEQ ID NO:2;
- 30 (ii) exhibits at least one of the following properties:
  - (a) ability to induce proliferation of vascular endothelial cells;
  - (b) ability to interact with *flt-1/flk-1* family of receptors;

- (c) ability to induce cell migration, cell survival and/or an increase in intracellular levels of alkaline phosphatase.

By "biologically isolated" is meant that the molecule has undergone at least one step of purification from a biological source. Preferably, the molecule is also biologically pure meaning that a composition comprises at least about 20%, more preferably at least about 40%, still more preferably at least about 65%, even still more preferably at least about 80-90% or greater of the molecule as determined by weight, activity or other convenient means, relative to other compounds in the composition. Most preferably, the molecule is sequenceably pure.

Another preferred aspect of the present invention provides the molecule in recombinant form.

According to this aspect of the present invention, there is provided a recombinant molecule comprising a sequence of amino acids which:

- (i) is at least about 15% similar to the amino acid sequence set forth in SEQ ID NO:2; and
- (ii) is at least 5% dissimilar to the amino acid sequence set forth in SEQ ID NO:2.

A related aspect of the present invention is directed to a recombinant molecule having the following characteristics:

- (i) comprises an amino acid sequence having at least about 15% similarity but at least about 5% dissimilarity to all or part of the amino acid sequence set forth in SEQ ID NO:2;
- (ii) exhibits at least one property in common with VEGF.

A further related aspect of the present invention contemplates a recombinant molecule having the following characteristics:

- (i) comprises an amino acid sequence having at least about 15% similarity but at least about 5% dissimilarity to the amino acid sequence set forth in SEQ ID NO:2;

- (ii) exhibits at least one of the following properties:
- (a) ability to induce proliferation of vascular endothelial cells;
  - (b) ability to interact with *flt-1/flk-1* family of receptors;
  - (c) ability to induce cell migration, cell survival and/or an increase in intracellular levels of alkaline phosphatase.

The present invention also contemplates genomic or partial genome clones encoding a proteinaceous molecule having at least about 15% amino acid similarity but at least about 5% dissimilarity to SEQ ID NO:1.

The amino acid sequence set forth in SEQ ID NO:2 corresponds to human VEGF (referred to herein as "VEGF<sub>165</sub>"). Accordingly, the molecule of the present invention is VEGF-like or is a homologue of VEGF but comprises an amino acid sequence which is similar but non-identical to the amino sequence of VEGF. Although the present invention is exemplified using a human VEGF-like molecule, this is done with the understanding that the instant invention contemplates the homologous molecule and encoding sequence from other mammals such as livestock animals (e.g. sheep, pigs, horses and cows), companion animals (e.g. dogs and cats) and laboratory test animals (e.g. mice, rats, rabbits and guinea pigs) as well as non-mammals such as birds (e.g. poultry birds), fish and reptiles. In a most preferred embodiment, the VEGF-like molecule is of human origin and encoded by a gene located at chromosome 11q13. The present invention extends, therefore, to the genomic sequence or part thereof encoding the subject VEGF-like molecule.

Preferably, the percentage similarity is at least about 30%, more preferably at least about 40%, still more preferably at least about 50%, still even more preferably at least about 60-70%, yet even more preferably at least about 80-95% to all or part of the amino acid sequence set forth in SEQ ID NO:2.

In a particularly preferred embodiment, the VEGF-like molecule of the present invention comprises a sequence of amino acids as set forth in SEQ ID NO:4 or is a part, fragment, derivative or analogue thereof. Particularly preferred similarities include about 19-20%,

and 29-30%. Reference herein to derivatives also includes splice variants. Accordingly, the present invention extends to splice variants of SOM175. Examples of splice variants contemplated by the present invention include but are not limited to variants with an amino acid sequence substantially as set forth in at least one of SEQ ID NO:6, SEQ ID  
5 NO:8 and/or SEQ ID NO:10 or mutants or derivatives or further splice variants thereof.

Another embodiment provides a recombinant molecule having the following characteristics:

- 10 (i) an amino acid sequence substantially as set forth in SEQ ID NO:4 or having at least about 15% similarity to all or part thereof provided that said amino acid sequence is at least about 5% dissimilar to all or part of the amino acid sequence set forth in SEQ ID NO:2;
- (ii) exhibits at least one biological property in common with VEGF.

15 Another embodiment provides a recombinant molecule having the following characteristics:

- 20 (i) an amino acid sequence substantially as set forth in SEQ ID NO:6 or having at least about 15% similarity to all or part thereof provided that said amino acid sequence is at least about 5% dissimilar to all or part of the amino acid sequence set forth in SEQ ID NO:2;
- (ii) exhibits at least one biological property in common with VEGF.

Another embodiment provides a recombinant molecule having the following characteristics:

- 25 (i) an amino acid sequence substantially as set forth in SEQ ID NO:8 or having at least about 15% similarity to all or part thereof provided that said amino acid sequence is at least about 5% dissimilar to all or part of the amino acid sequence set forth in SEQ ID NO:2;
- (ii) exhibits at least one biological property in common with VEGF.

30

Another embodiment provides a recombinant molecule having the following characteristics:

- (i) an amino acid sequence substantially as set forth in SEQ ID NO:10 or having at least about 15% similarity to all or part thereof provided that said amino acid sequence is at least about 5% dissimilar to all or part of the amino acid sequence set forth in SEQ ID NO:2;
- 5 (ii) exhibits at least one biological property in common with VEGF.

Such properties of VEGF include at least one of:

- (a) ability to induce proliferation of vascular endothelial cells;
- (b) an ability to interact with *flt-1/flk-1* family of receptors;
- 10 (c) an ability to induce cell migration, cell survival and/or an increase in intracellular levels of alkaline phosphatase.

In accordance with the present invention, a preferred similarity is at least about 40%, more preferably at least about 50% and even more preferably at least about 65%  
15 similarity.

Still a further aspect of the present invention contemplates a peptide fragment corresponding to a portion of the amino acid sequence set forth in SEQ ID NO:4 or a splice variant thereof such as set forth in SEQ ID NO:6, SEQ ID NO:8 or SEQ ID  
20 NO:10 or a chemical equivalent thereof. The biologically isolated or recombinant molecule of the present invention may be naturally glycosylated or may comprise an altered glycosylation pattern depending on the cells from which it is isolated or synthesised. For example, if produced by recombinant means in prokaryotic organisms, the molecule would be non-glycosylated. The molecule may be a full length, naturally  
25 occurring form or may be a truncated or otherwise derivatised form.

Yet another aspect of the present invention is directed to a nucleic acid molecule encoding the VEGF-like molecule herein described. More particularly, the present invention provides a nucleic acid molecule comprising a sequence of nucleotides  
30 substantially as set forth in SEQ ID NO:3 or having at least 15% similarity to all or part thereof or being capable of hybridising under low stringency conditions to a reverse complement of the nucleotide sequence as set forth in SEQ ID NO:3 provided that the

nucleic acid sequence having at least 15% similarity but at least 30% dissimilarity to the nucleotide sequence as set forth in SEQ ID NO:3. The nucleotide sequence set forth in SEQ ID NO:3 is also referred to herein as "SOM175". Preferably, the percentage dissimilarity is about 35%, more preferably about 39% and even more preferably about 40-50% or greater.

For the purposes of defining the level of stringency, reference can conveniently be made to Sambrook *et al* (1989) at pages 9.47-9.51 which is herein incorporated by reference where the washing steps disclosed are considered high stringency. A low stringency is defined herein as being in 4-6X SSC/0.1-0.5% w/v SDS at 37-45°C for 2-3 hours. Depending on the source and concentration of nucleic acid involved in the hybridisation, alternative conditions of stringency may be employed such as medium stringent conditions which are considered herein to be 1-4X SSC/0.25-0.5% w/v SDS at  $\geq 45^{\circ}\text{C}$  for 2-3 hours or high stringent conditions considered herein to be 0.1-1X SSC/0.1% w/v SDS at 60°C for 1-3 hours.

The present invention further contemplates a nucleic acid molecule which encodes a VEGF-like molecule as hereinbefore described having at least 15% nucleotide sequence homology to SEQ ID NO:3. Preferred levels of homology include at least about 40%, more preferably around 60-70%.

The present invention is further directed to the murine homologue of human VEGF (referred to herein as "mVRF"). The mVRF has approximately 85% identity and 92% conservation of amino acid residues over the entire coding region compared to human VEGF. The mVRF is encoded by a nucleic acid molecule comprising a nucleotide sequence substantially as set forth in Figure 9.

The VEGF-like molecule of the present invention will be useful in the development of a range of therapeutic and/or diagnostic applications alone or in combination with other molecules such as VEGF. The present invention extends, therefore, to pharmaceutical compositions comprising the VEGF-like molecule or parts, fragments, derivatives, homologues or analogues thereof together with one or more pharmaceutically acceptable

carriers and/or diluents. Furthermore, the present invention extends to vectors comprising the nucleic acid sequence set forth in SEQ ID NO:3 or having at least about 15%, more preferably about 40% and even more preferably around 60-70% similarity thereto but at least 30% and more preferably around 39% dissimilarity thereto and host  
5 cells comprising same. In addition, the present invention extends to ribozymes and antisense molecules based on SEQ ID NO:3 as well as neutralizing antibodies to the VEGF-like molecule. Such molecules may be useful in ameliorating the effects of, for example, over expression of VEGF-like genes leading to angiogenesis or vascularization of tumours.

10

Another aspect of the present invention contemplates a method of inducing astroglial proliferation in a mammal, said method comprising administering to said mammal an effective amount of a recombinant proteinaceous molecule having the characteristics:

- 15 (i) comprises an amino acid sequence having at least about 15% similarity but at least about 5% dissimilarity to the sequence set forth in SEQ ID NO:2;
- (ii) exhibits at least one property in common with vascular endothelial growth factor (VEGF),

said administration being for a time and under conditions sufficient to induce astroglial  
20 proliferation.

Preferably, the recombinant proteinaceous molecule comprises the amino acid sequence set forth in SEQ ID NO:3 or SEQ ID NO:6.

25 A further aspect of the present invention provides a method of promoting neural survival and/or proliferation in a mammal, said method comprising administering to said mammal an effective amount of a recombinant proteinaceous molecule having the characteristics:

- 30 (i) comprises an amino acid sequence having at least about 15% similarity but at least about 5% dissimilarity to the sequence set forth in SEQ ID NO:2;
- (ii) exhibits at least one property in common with vascular endothelial growth factor (VEGF),



- 9 -

said administration being for a time and under conditions sufficient to induce astroglial proliferation.

Preferably, the recombinant proteinaceous molecule comprises the amino acid sequence  
5 set forth in SEQ ID NO:3 or SEQ ID NO:6.

The present invention also contemplates antibodies to the VEGF-like molecule or nucleic acid probes to a gene encoding the VEGF-like molecule which are useful as diagnostic agents.

10

The present invention is further described by reference to the following non-limiting Figures and/or Examples.

In the Figures:

15

**Figure 1** Nucleotide sequence [SEQ ID NO:1] and corresponding amino acid sequence [SEQ ID NO:2] of VEGF<sub>165</sub>.

**Figure 2** Nucleotide sequence [SEQ ID NO:3] and corresponding amino acid  
20 sequence [SEQ ID NO:4] of SOM175.

**Figure 3** Results of BLAST search with SOM175 protein sequence.

**Figure 4** BESTFIT alignment of VEGF cDNA and SOM175 cDNA.

25

**Figure 5** Multiple alignment of VEGF<sub>165</sub> with SOM175 and its splice variants at the nucleotide level.

**Figure 6** Multiple alignment of VEGF<sub>165</sub> with SOM175 and its splice variants at  
30 the amino acid level.

**Figure 7** Diagrammatic representation of SOM175 and its splice variants.

- 10 -

**Figure 8(a)** Diagrammatic representation of genomic structure of human SOM175 genomic showing exon/intron map.

**Figure 8(b)** Diagrammatic representation of genomic structure of human SOM175 showing exon/intron boundaries.

**Figure 9** Nucleotide and predicted peptide sequences derived from mVRF cDNA clones. Numbering of nucleotides are given on the left, starting from the A of the initiation codon. Amino acids are numbered on the right, starting from the first residue of the predicted mature protein after the putative signal peptide has been removed. The alternately spliced region is double underlined and the resulting peptide sequence from each mRNA is included. A potential polyadenylation signal is indicated in boldface. Start and stop codons of mVRF<sub>167</sub> and mVRF<sub>186</sub> are underlined and a polymorphic AC repeat in the 3' UTR is indicated by a stippled box. The positions of intron/exons boundaries are indicated by arrowheads.

**Figure 10** BESTFIT alignments of human and murine VRF protein isoforms. A: mVRF<sub>167</sub> and hVRF<sub>167</sub>. B: mVRF<sub>186</sub> and hVRF<sub>186</sub> from the point where the sequences diverge from the respective 167 amino acid isoforms. Amino acid identities are marked with vertical bars and conserved amino acids with colons. An arrow marks the predicted signal peptide cleavage site of human and mouse VRF.

**Figure 11** BESTFIT alignment of mVRF<sub>167</sub> and mVEGF<sub>188</sub> (Breier *et al*, 1992) peptide sequences. An arrow marks the signal peptide cleavage site of mVEGF. Identical amino acids are indicated by vertical bars and conservative substitutions by colons. Numbering of amino acids is as described in the legend to Figure 9.

**Figure 12** Comparison of gene structure between VRF (a generic VRF gene is shown since the intron/exon organisation of the mouse and human homologues is almost identical) and other members of the human VEGF/PIGF/PDGF gene family. Exons are represented by boxes. Protein coding regions and untranslated regions are shown by filled and open sections respectively. The hatched region in VRF indicates the

additional 3' UTR sequence formed by alternate splicing of the VRF<sub>186</sub> isoform. Potential alternate splice products of each gene are shown.

**Figure 13** Autoradiogram of a Northern blot of total RNA from various adult mouse tissues (as indicated) hybridised with an mVRF cDNA clone. A major transcript of 1.3 kb was detected in all samples.

**Figure 14** Film autoradiographs (A-C) and dark-field micrographs (D-E) illustrating the expression pattern of mVRF and mRNA in the mouse. In the E14 mouse embryo (A) positive signals are present over the developing heart (Ha) and cerebral cortex (Cx). A low background signal is also present over other tissues in the section. In the E17 embryo (B) and the heart (Ha) is clearly visible due to a strong hybridisation signal. An equally strong signal is present over brown adipose tissue (Fa) in the back and around the thoracic cage. A moderate hybridisation signal is present over the spinal cord (SC) and the tongue (T). The background signal is reduced compared with the E14 embryo. In the young adult mouse (C-D), positive signals are present over the heart (Ha) and adipose tissue (Fa) around the thoracic cage, while, for example, the lungs (Lu) are unlabeled). The hybridisation signal over the heart is evenly distributed over the entire left ventricle, including papillary muscles (D). In the E17 heart hybridised with an excess of cold probe, no positive signal is present (E). Scale bars = 0.5 mm (A), 1.2 mm (B), 1 mm (C), 0.3 mm (D), 0.1 mm (E).

**Figure 15** Dark - (A and C) and bright-field (B and D) micrographs showing mVRF mRNA expression in mouse adipose tissue (A-B) and spinal cord (C-D). A strong hybridisation signal is present over fat (A), as shown by the strong labeling in Sudan black stained sections (B). A weak signal is present also in skeletal muscle (M in A-B). In the adult spinal cord (C) the mVRF probes gave a neuronal staining pattern over the gray matter. Toluidine counterstaining showing that motoneurons in the ventral horn (D), interneurons in the deep part of the dorsal horn and around the central canal (not shown) were largely positive for mVRF mRNA. Scale bars = 0.1 mm (A), 0.1 mm (B), 0.25 mm (C), 0.015 mm (D).

**Figure 16** Effect of VEGF on embryonic day 8 (E8) chick sensory neurons as determined by % survival, % neurite outgrowth and average neurite length ( $\mu\text{m}$ ).

**Figure 17** Effects of VEGF and SOM175 on chick glia. Tested were CNS glial,  
5 peripheral glia and CNS oligodendrocytes.

**Figure 18** Effect of various SOM175 proteins on mouse astroglial cells. ■  $^3\text{H}$  (cpm)

1. FGF-2 (10 ng/ml) positive control
2.  $\text{SOM}\Delta\text{X6}^*$  1 ng/ml
- 10 3.  $\text{SOM}\Delta\text{X6}$  10 ng/ml
4.  $\text{SOM}\Delta\text{X6}$  100 ng/ml
5.  $\text{SOM}\Delta\text{X6}$  1000 ng/ml
6.  $\text{SOM}\Delta\text{X6}$  1000 ng/ml, no heparin
7.  $\text{SOMX6}^{**}$  1 ng/ml
- 15 8.  $\text{SOMX6}$  10 ng/ml
9.  $\text{SOMX6}$  100 ng/ml
10.  $\text{SOMX6}$  1000 ng/ml
11.  $\text{SOMX6}$  1000 ng/ml, no heparin

\* This refers to SOM175 absent exon 6;

20 \*\* This refers to SOM175.

**Figure 19** Effect of various SOM175 proteins on mouse oligodendroglial cells. ■  $^3\text{H}$  (cpm)

1. FGF-2 (10 ng/ml) positive control
- 25 2.  $\text{SOM}\Delta\text{X6}^*$  1 ng/ml
3.  $\text{SOM}\Delta\text{X6}$  10 ng/ml
4.  $\text{SOM}\Delta\text{X6}$  100 ng/ml
5.  $\text{SOM}\Delta\text{X6}$  1000 ng/ml
6.  $\text{SOM}\Delta\text{X6}$  1000 ng/ml, no heparin
- 30 7.  $\text{SOMX6}^{**}$  1 ng/ml
8.  $\text{SOMX6}$  10 ng/ml
9.  $\text{SOMX6}$  100 ng/ml

10. SOMX6 1000 ng/ml
11. SOMX6 1000 ng/ml, no heparin

\* This refers to SOM175 absent exon 6;

\*\* This refers to SOM175.

5

**Figure 20** Effect of various SOM175 proteins on mouse forebrain neurons. ■ % survival

1. FGF-2 (10 ng/ml) positive control
2. SOM $\Delta$ X6\* 1 ng/ml
- 10 3. SOM $\Delta$ X6 10 ng/ml
4. SOM $\Delta$ X6 100 ng/ml
5. SOM $\Delta$ X6 1000 ng/ml
6. SOM $\Delta$ X6 1000 ng/ml, no heparin
7. SOMX6\*\* 1 ng/ml
- 15 8. SOMX6 10 ng/ml
9. SOMX6 100 ng/ml
10. SOMX6 1000 ng/ml
11. SOMX6 1000 ng/ml, no heparin

\* This refers to SOM175 absent exon 6;

20 \*\* This refers to SOM175.

TABLE 1  
SUMMARY OF SEQUENCE IDENTITY NUMBERS

---

5	SEQ ID NO:1	Nucleotide sequence of VEGF <sub>165</sub>
	SEQ ID NO:2	Amino acid sequence of VEGF <sub>165</sub>
	SEQ ID NO:3	Nucleotide sequence of SOM175 (VEGF-like molecules)
	SEQ ID NO:4	Amino acid sequence of SOM175
10	SEQ ID NO:5	Nucleotide sequence of SOM175 absent exon 6
	SEQ ID NO:6	Amino acid sequence of SOM175 absent exon 6
	SEQ ID NO:7	Nucleotide sequence of SOM175 absent exon 6 and exon 7
	SEQ ID NO:8	Amino acid sequence of SOM175 absent exon 6 and exon 7
	SEQ ID NO:9	Nucleotide sequence of SOM175 absent exon 4
15	SEQ ID NO:10	Amino acid sequence of SOM175 absent exon 4
	SEQ ID NO:11	Oligonucleotide
	SEQ ID NO:12	Oligonucleotide
	SEQ ID NO:13	Oligonucleotide
	SEQ ID NO:14	Oligonucleotide

---

20

### EXAMPLE 1

#### Human cDNA clones

The original SOM175 cDNA was isolated by screening a human foetal brain library  
 25 (λzapII, Stratagene) with the cosmid D11S750 (Larsson *et al*, 1992). The plasmid was  
 excised "*in vivo*" and a single 1.1kb cDNA was obtained. Three independent SOM175  
 cDNAs clones were also isolated from a human foetal spleen library (Stratagene, Uni-  
 zap) using the above-mentioned SOM175 insert as a probe. Three clones were obtained:  
 SOM175-4A, -5A and -6A. SOM175-5A is an alternately spliced clone with exon 4  
 30 being absent (SOM175-e4). These library screens were performed using hybridisation  
 conditions recommended by the manufacturer of the library (Stratagene) and random  
 primed insert of SOM175.

Two partial human SOM175 cDNAs have also isolated from a  $\lambda$ GT11 human melanoma cell line A2058 library (Clontech) cDNA library screens were performed using hybridisation conditions described by Church and Gilbert, 1984). In each case, the probe was generated by random priming of a PCR product derived from SOM175 (18f-  
5 700r).

#### Mouse cDNA Clones

Human SOM175 was also used to screen a mouse neonatal whole brain cDNA library (Unizap, Stratagene). Four non-chimeric clones were isolated: M175-A, B, C, D. All  
10 clones were partial cDNAs and M175-C contained several introns. Three of these cDNAs lacked the exon 6.

Another clone referred to as M1 was completely sequenced and was found to contain the full open reading frame plus part of the 5'utr and total 3'utr.

15

### EXAMPLE 2

#### DNA SEQUENCE ANALYSIS

The entire sequence of the cDNA clone (SOM175) was compiled and is shown in Figure 2 with its corresponding amino acid sequence. This sequence was screened for  
20 open reading frames using the MAP program (GCG, University of Wisconsin). A single open reading frame of 672bp was observed (see Figure 2). There appears to be little 5' untranslated sequences (2bp). The 3' untranslated region appears to be complete as it includes a poly-adenylation signal and poly-A tail.

25 Database homology searches were performed using the BLAST algorithm (run at NCBI, USA). This analysis revealed homology to several mammalian forms of VEGF (see Figure 3). The amount of homology between SOM175 and human VEGF<sub>165</sub> was determined using the BESTFIT program (GCG, University of Wisconsin; see Figures 4 and 5). Nucleotide homology was estimated at 69.7% and protein homology was  
30 estimated as at least 33.3% identity and 52.5% conservation using BESTFIT analysis. BLAST analysis on nucleotide sequences revealed the almost complete match to a human expressed sequence tag EST06302 (Adams *et al.*, 1993).

These data indicate that SOM175 encodes a growth factor that has structural similarities to VEGF. Both genes show start and stop codons in similar positions and share discrete blocks of homology. All 8 cysteines as well as a number of other VEGF residues believed to be involved in dimerisation are conserved. These residues are Cysteine-47, 5 Proline-70, Cysteine-72, Valine-74, Arginine-77, Cysteine-78, Glycine-80, Cysteine-81, Cysteine-82, Cysteine-89, Proline-91, Cysteine-122 and Cysteine-124 and are shown in Figure 6. Given the structural conservation between VEGF and the SOM175 gene product it is also possible that they share functional similarities. It is proposed that SOM175 encodes a VEGF-like molecule that shares some properties with VEGF but has 10 unique properties of its own. The nucleotide sequence and corresponding amino acid sequence of VEGF<sub>165</sub> is shown in Figure 1.

### EXAMPLE 3

The percentage similarity and divergence between VEGF<sub>165</sub> family and SOM175 family 15 (protein) were analysed using the Clustal method, MegAlign Software, DNASTAR, Wisconsin. The results are shown in Tables 2.1 and 2.2. The alternatively spliced forms of SOM175 are abbreviated to SOM715-e6 where all of exon 6 is deleted; SOM715-e6 and 7 where all of exons 6 and 7 are deleted; and SOM175-e4 where all of exon 4 is deleted. The spliced form of SOM175 are shown in Figure 7. Genomic 20 maps of SOM175 showing intron/exon boundaries are shown in Figure 8a and 8b.



Table 2.1

A Percent nucleotide similarity between splice variants of SOM175 and  
5 human VEGF<sub>165</sub>

	VEGF <sub>165</sub>	SOM175	SOM175-e6	SOM175-e6&7	SOM175-e4
VEGF <sub>165</sub>	***	34.9	39.7	41.4	37.0
10 SOM175		***	98.9	95.1	99.2
SOM175-e6			***	98.8	84.0
SOM175-e6&7				***	80.3
SOM175-e4					***

15

- 18 -

**B Percent nucleotide divergence between splice variants of SOM175 and human VEGF<sub>165</sub>**

5		VEGF <sub>165</sub>	SOM175	SOM175-e6	SOM175-e6&7	SOM175-e4
	VEGF <sub>165</sub>	***	41.7	41.6	41.7	41.8
	SOM175		***	0.2	0.2	0.0
	SOM175-e6			***	0.0	0.2
10	SOM175-e6&7				***	0.3
	SOM175-e4					***

Table 2.2

**15 A Percent amino acid identity between splice variants of SOM175 and human VEGF<sub>165</sub>**

		VEGF <sub>165</sub>	SOM175	SOM175-e6	SOM175-e6&7	SOM175-e4
20	VEGF <sub>165</sub>	***	31.4	42.3	33.5	40.6
	SOM175		***	74.7	73.7	99.1
	SOM175-e6			***	76.8	99.1
	SOM175-e6&7				***	99.1
	SOM175-e4					***

25

**B Percent amino acid divergence between splice variants of SOM175 and human VEGF<sub>165</sub>**

	VEGF <sub>165</sub>	SOM175	SOM175-e6	SOM175-e6&7	SOM175-e4
VEGF <sub>165</sub>	***	65.7	55.4	54.6	57.4
SOM175		***	19.9	4.2	0.0
SOM175-e6			***	0.0	0.0
SOM175-e6&7				***	0.0
SOM175-e4					***

15

**EXAMPLE 4**

**BIOASSAYS TO DETERMINE THE FUNCTION OF SOM175**

Assays are conducted to evaluate whether SOM175 has similar activities to VEGF on endothelial cell function, angiogenesis and wound healing. Other assays are performed based on the results of receptor binding distribution studies.

20

**Assays of endothelial cell function**

*Endothelial cell proliferation.* Endothelial cell growth assays as described in Ferrara & Henzel (1989) and in Gospodarowicz *et al* (1989).

25 *Vascular permeability assay.* This assay, which utilises the Miles test in guinea pigs, will be performed as described in Miles & Miles (1952).

*Cell adhesion assay.* The influence of SOM175 on adhesion of polymorphs to endothelial cells is analysed.

30

*Chemotaxis.* This is performed using the standard Boyden chamber chemotaxis assay.

*Plasminogen activator assay.* Endothelial cells are tested for plasminogen activator and plasminogen activator inhibitor production upon addition of SOM175 (Pepper *et al* (1991)).

- 5    *Endothelial cell migration assay.* The ability of SOM175 to stimulate endothelial cells to migrate and form tubes is assayed as described in Montesano *et al* (1986).

#### **Angiogenesis Assay**

- SOM175 induction of an angiogenic response in chick chorioallantoic membrane is  
10    evaluated as described in Leung *et al* (1989).

Possible neurotrophic actions of SOM175 are assessed using the following assays:

#### **Neurite outgrowth assay and gene induction (PC12 cells)**

- 15    PC12 cells (a pheochromocytoma cell line) respond to NGF and other neurotrophic factors by developing the characteristics of sympathetic neurons, including the induction of early and late genes and the extension of neurites. These cells are exposed to SOM175 and their response monitored (Drinkwater *et al* (1991); and Drinkwater *et al* (1993)).

20

#### **Cultured neurons from the Peripheral Nervous System (PNS)**

Primary cultures of the following PNS neurons are exposed to SOM175 and monitored for any response:

- sensory neurons from neural crest and dorsal root ganglia
- 25    - sympathetic neurons from sympathetic chain ganglia
- placode derived sensory neurons from nodose ganglia
- motoneurons from spinal cord

The assays are described in Suter *et al* (1992) and in Marinou *et al* (1992).

- 30    Where an *in vitro* response is observed, *in vivo* assays for properties such as uptake and retrograde transport are performed as described in Hendry *et al* (1992).

### **Nerve regeneration (PNS)**

Where neurotrophic effects of SOM175 are observed, its possible role in the regeneration of axotomised sensory neurons, sympathetic neurons and motoneurons is analysed by the methods of Otto *et al* (1989); Yip *et al* (1984) and Hendry *et al* 5 (1976).

### **Actions of SOM175 on CNS neurons**

The ability of SOM175 to promote survival of central nervous system neurons is analysed as described in Hagg *et al* (1992); Williams *et al* (1986); Hefti (1986) and 10 Kromer (1987).

### **Wound Healing**

The ability of SOM175 to support wound healing are tested in the most clinically relevant model available, as described in Schilling *et al* (1959) and utilised by Hunt 15 *et al* (1967).

### **The Haemopoietic System**

A variety of *in vitro* and *in vivo* assays on specific cell populations of the haemopoietic system are available and are outlined below:

#### 20 ***Stem Cells***

##### ***Murine***

A variety of novel *in vitro* murine stem cell assays have been developed using FACS-purified cells:

#### 25 (a) **Repopulating Stem Cells**

These are cells capable of repopulating the bone marrow of lethally irradiated mice, and have the Lin<sup>-</sup>, Rh<sup>hi</sup>, Ly-6A/E<sup>+</sup>, c-kit<sup>+</sup> phenotype. The test substance is tested on these cells either alone, or by co-incubation with multiple factors, followed by measurement of cellular proliferation by <sup>3</sup>H thymidine incorporation.

**(b) Late Stage Stem Cells**

These are cells that have comparatively little bone marrow repopulating ability but can generate D13 CFU-S. These cells have the Lin<sup>-</sup>, Rh<sup>hi</sup>, Ly-6A/E<sup>+</sup>, c-kit<sup>+</sup> phenotype. The test substance is incubated with these cells for a period of time, injected into lethally irradiated recipients, and the number of D13 spleen colonies enumerated.

**(c) Progenitor-Enriched Cells**

These are cells that respond *in vitro* to single growth factors, and have the Lin<sup>-</sup>, Rh<sup>hi</sup>, Ly-6A/E<sup>+</sup>, c-kit<sup>-</sup> phenotype. This assay will show if SOM175 can act directly on haemopoietic progenitor cells. The test substance is incubated with these cells in agar cultures, and the number of colonies enumerated after 7-14 days.

**15 Atherosclerosis**

Smooth muscle cells play a crucial role in the development or initiation of atherosclerosis, requiring a change in their phenotype from a contractile to a synthetic state. Macrophages, endothelial cells, T lymphocytes and platelets all play a role in the development of atherosclerotic plaques by influencing the growth and phenotypic modulations of smooth muscle cell. An *in vitro* assay that measures the proliferative rate and phenotypic modulations of smooth muscle cells in a multicellular environment is used to assess the effect of SOM175 on smooth muscle cells. The system uses a modified Rose chamber in which different cell types are seeded onto opposite coverslips.

25

**Effects of SOM175 on bone**

The ability of SOM175 to regulate proliferation of osteoblasts is assayed as described in Lowe *et al* (1991). Any effects on bone resorption are assayed as described in Lowe *et al* (1991). Effects on osteoblast migration and changes in intracellular molecules (e.g. cAMP accumulation, alkaline phosphatase levels) are analysed as described in Midy *et al* (1994).

30

### Effects on skeletal muscle cells

Effects of SOM175 on proliferation of myoblasts and development of myotubes can be determined as described by Ewton *et al* (1980) and by Gospodarowicz *et al* (1976).

5

## EXAMPLE 5

### CLONING MURINE VEGF DNA

#### Isolation of cDNAs

Murine VRF (mVRF) clones were selected from a lambda Zap new born whole brain  
10 cDNA library (Stratagene). Primary phage from high density filters ( $5 \times 10^4$   
pfu/plate) were identified by hybridisation with a 682bp  $^{32}\text{P}$ -labelled probe generated  
by PCR from an hVRF cDNA (pSOM175) as described above. Hybridisation and  
stringent washes of nylon membranes (Hybond-N) were carried out at  $65^\circ\text{C}$  under  
conditions described by Church and Gilbert (1984). Positive plaques were picked,  
15 purified and excised *in vivo* to produce bacterial colonies containing cDNA clones in  
pBluescript SK-.

#### Isolation of genomic clones

Genomic clones were isolated from a mouse strain SV/129 library cloned in the  
20 lambda Fix II vector (Stratagene). High density filters ( $5 \times 10^4$  pfu/filter) were  
screened with a 563 bp  $^{32}\text{P}$ -labelled probe generated by PCR amplification of the  
nucleotide 233-798 region of the mVRF cDNA (see Figure 9). Positive clones were  
plugged and re-screened with filters containing 400-800 pfu. Large scale phage  
preparations were prepared using the QIAGEN lambda kit or by  $\text{ZnCl}_2$  purification  
25 (Santos, 1991).

#### Nucleotide sequencing and analysis

cDNAs were sequenced on both strands using a variety of vector-based and internal  
primers with Applied Biosystems Incorporated (ABI) dye terminator sequencing kits  
30 according to the manufacturer's specifications. Sequences were analysed on an ABI  
Model 373A automated DNA sequencer. Peptide homology alignments were  
performed using the program BESTFIT (GCG, Wisconsin).

### Identification of intron/exon boundaries

Identification of exon boundaries and flanking regions was carried out using PCR with mouse genomic DNA or mVRF genomic lambda clones as templates. The primers used in PCR to identify introns were derived from the hVRF sequence and to  
5 allow for potential human-mouse sequence mismatches annealing temperatures 5-10°C below the estimated  $T_m$  were used. All PCR products were sized by agarose gel electrophoresis and gel purified using QIAquick spin columns (Qiagen) and the intron/exon boundaries were sequenced directly from these products. In addition, some splice junctions were sequenced from subcloned genomic fragments of MVRF.  
10 Intron/exon boundaries were identified by comparing cDNA and genomic DNA sequences.

### Northern analysis

Total cellular RNA was prepared from a panel of fresh normal adult mouse tissues  
15 (brain, kidney, liver, muscle) using the method of Chomczynski and Sacchi (1987). 20µg of total RNA were electrophoresed, transferred to a nylon membrane (Hybond N, Amersham) and hybridised under standard conditions (Church & Gilbert, 1984). Filters were washed at 65°C in 0.1xSSC (20xSSC is 3M NaCl/0.3M trisodium citrate), 0.1% SDS and exposed to X-ray film with intensifying screens at -70°C for  
20 1-3 days.

### Characterisation of mVRF cDNAs

Murine VRF homologues were isolated by screening a murine cDNA library with an hVRF cDNA clone. Five clones of sizes varying from 0.8-1.5 kb were recovered  
25 and sequenced. The cDNA sequences were compiled to give a full length 1041 bp cDNA sequence covering the entire open reading frame (621 bp or 564 bp depending on the splice form, see below) and 3' UTR (379 bp), as well as 163 bp of the 5' UTR (Figure 9).

30 The predicted initiation codon matched the position of the start codon in hVRF. One other out of frame ATG was located at position -47 and two termination codons were observed upstream (positions -9 and -33, respectively) and in-frame with the putative



initiation codon.

The predicted N-terminal signal peptide of hVRF appears to be present in mVRF with 81% identity (17/21 amino acids). Peptide cleavage within mVRF is expected to occur after residue 21 (Figure 10). These data suggest that mature mVRF is secreted and could therefore conceivably function as a growth factor.

As with hVRF, two open reading frames (ORFs) were detected in cDNAs isolated by library screening. Four of five clones were found to be alternatively spliced and lacked a 101 bp fragment homologous to exon 6 of hVRF. The predicted peptide sequences of the two isoforms of mVRF were determined and aligned with the corresponding human isoforms (Figure 10).

The message encoding mVRF<sub>186</sub> contains a 621 bp ORF with coding sequences terminating at position +622, towards the end of exon 7 (Figure 9). The smaller message encoding mVRF<sub>167</sub> actually terminates downstream of the +622 TAG site due to a frame shift resulting from splicing out of the 101 bp exon 6 and the introduction of a stop codon (TGA) at position +666, near the beginning of exon 8 (Figure 9).

The mVRF<sub>186</sub> protein has strong homology to the amino and central portions of VEGF while the carboxyl end is completely divergent and is alanine rich. mVRF<sub>167</sub> possesses these similarities and also maintains homology to mVEGF right through to the C-terminus (Figure 11). The overall homology of mVRF<sub>167</sub> to hVRF<sub>167</sub> was 85% identity and 92% similarity, respectively (Figure 10). Likewise, homology between mVRF<sub>167</sub> and mVEGF (Breier *et al*, 1992) was 49% identity and 71% conservative amino acid substitution, respectively (Figure 11).

A canonical vertebrate polyadenylation signal (AATAAA) (Birnstiel *et al*, 1986) was not present in the mVRF cDNA, however, the closely matching sequence GATAAA is present at similar positions in both mouse and human VRF cDNAs (Figure 9). In contrast to hVRF, mVRF was found to contain an AC dinucleotide repeat at the

extreme 3' end of the 3' UTR (nucleotide positions 998 to 1011, Figure 9). Polymorphism of this repeat region was observed between some of the mVRF cDNAs, with the number of dinucleotides varying from 7 to 11.

## 5 Genomic characterisation of mVRF

Intron/exon boundaries (Table 3) were mapped using primers which flanked sequences homologous to the corresponding hVRF boundaries. Introns I, III, IV and VI of mVRF (Table 3, Figure 12) were smaller than the hVRF intervening sequences. The complete genomic sequence was compiled from the 5' UTR of  
10 mVRF through to intron VI, the largest intervening region (2.2 kb), by sequencing amplified introns and cloned genomic portions of mVRF. There was only one major difference in genomic structure between mVRF and hVRF and that was the exon 7/intron VI boundary of mVRF was located 10bp further downstream in relation to the cDNA sequence, hence exon 7 in mVRF is 10bp longer than the corresponding  
15 exon in hVRF.

Exons 6 and 7 are contiguous in mVRF, as has been found to occur in the human homologue. The strong sequence homology between exon 6 of mVRF and hVRF (Figure 10) suggests that this sequence is not a retained intronic sequence but rather  
20 encodes a functional part of the VRF<sub>186</sub> isoform.

General intron/exon structure is conserved between the various members of the VEGF gene family (VEGF, PIGF, hVRF) and therefore it is not surprising that the overall genomic organisation of the mVRF gene is very similar to these genes  
25 (Figure 12).

Previous comparative mapping studies have shown that the region surrounding the human multiple endocrine neoplasia type 1 disease locus on chromosome 11q13 is syntenic with the proximal segment of mouse chromosome 19 (Rochelle *et al*, 1992).  
30 Since the inventors have mapped the hVRF gene to within 1kb of the human *MEN1* locus (see above) it is most likely that the murine VRF gene maps near the centromere of chromosome 19.

### Expression studies of mVRF

Northern analysis of RNA from adult mouse tissues (muscle, heart, lung and liver) showed that expression appears to be ubiquitous and occurs primarily as a major band of approximately 1.3kb in size (Figure 14). This is somewhat different to the pattern observed for hVRF in which two major bands of 2.0 and 5.5 kb have been identified in all tissues examined. The 1.3 kb murine message presumably corresponds to the shorter of the human transcripts and the size variation thereof is most likely due to a difference in the length of the respective 5' UTRs.

### EXAMPLE 6

#### EXPRESSION OF MURINE VEGF IN PRE- AND POST-NATAL MOUSE

##### Animals

Timed pregnant (n=4) and young adult (n=2) mice (C57 inbred strain, ALAB, Sweden) were sacrificed with carbon dioxide, and the relevant tissues were taken out and frozen on a chuck. Tissues were kept at -70°C until further use. Two gestational ages were used in this study; embryonic day 8 (E8), 14 and E17.

##### *In situ* hybridisation histochemistry

*In situ* hybridisation was performed as previously described (Dagerlind *et al*, 1992).

Briefly, transverse sections (14µm) were cut in a cryostat (Microm, Germany), thawed onto Probe-On slides (Fisher Scientific, USA) and stored in black sealed boxes at -70°C until used. The sequences of the synthetic 42-mer oligonucleotides complementary to mRNA encoding mVRF were

ACCACCACCTCCCTGGGCTGGCATGTGGCACGTGCATAAACG [SEQ ID

NO:11] (complementary to nt 120-161) and

AGTTGTTTGACCACATTGCCCATGAGTTCCATGCTCAGAGGC [SEQ ID

NO:12] (complementary to nt 162-203). To detect the two alternative splice forms oligonucleotide GATCCTGGGGCTGGAGTGGGATGGATGATGTCAGCTGG [SEQ

ID NO:13] (complementary to nt xxx-xxx) and

GCGGGCAGAGGATCCTGGGGCTGTCTGGCCTCACAGCACT [SEQ ID NO:14]

were used. The probes were labeled at the 3'-end with deoxyadenosine-alpha[thio]triphosphate [<sup>35</sup>S] (NEN, USA) using terminal deoxynucleotidyl

transferase (IBI, USA) to a specific activity of  $7-10 \times 10^8$  cpm/ $\mu$ g and hybridised to the sections without pretreatment for 16-18 h at 42°C. The hybridisation mixture contained: 50% v/v formamide, 4 x SSC (1 x SSC = 0.15M NaCl and 0.015M sodium-citrate), 1 x Denhardt's solution (0.02% each of polyvinyl-pyrrolidone, BSA and Ficoll), 1% v/v sarcosyl (N-lauroylsarcosine; Sigma), 0.02M phosphate buffer (pH 7.0), 10% w/v dextran sulfate (Pharmacia, Sweden), 250 $\mu$ g/ml yeast tRNA (Sigma), 500 $\mu$ g/ml sheared and heat denatured salmon sperm DNA (Sigma) and 200mM dithiothreitol (DTT; LKB, Sweden). In control sections, the specificity of both probes was checked by adding a 20-fold excess of unlabeled probe to the hybridisation mixture. In addition, adjacent sections were hybridised with a probe unrelated to this study which gave a different expression pattern. Following hybridisation the sections were washed several times in 1 x SSC at 55°C, dehydrated in ethanol and dipped in NTB2 nuclear track emulsion (Kodak, USA). After 3-5 weeks the sections were development in D-19 developer (Kodak, USA) and cover-slipped. In some cases, sections were opposed to an autoradiographic film (Beta-max autoradiography film Amersham Ltd, UK) prior to emulsion-dipping.

The four different probes gave identical hybridisation patterns in all tissues examined. Mouse VRF expression was detecting already in the E8 embryo, in which positive signal was recorded over structures most likely corresponding to the neuronal tube. In sagittal sections of E14 mouse embryo the strongest hybridisation signal was present over heart and in the nervous system, especially cerebral cortex (Figure 14A). A low level of expression was present in all other tissues. At a later gestational age, E17, a high mVRF mRNA signal was confined to the heart and brown fat tissue in the back and around the neck (Figure 14B). Clearly positive hybridisation signals were present in the gray of the spinal cord and in the tongue (Figure 14B). Expression in the cerebral cortex was clearly reduced compared to day 14. The weak background expression seen in the E14 embryo in for example muscle, had decreased at this gestational age. A strong mVRF mRNA hybridisation signal was present solely over the heart and in the brown fat in the young adult mice (Figure 14C). The signal over the heart was evenly distributed over the entire ventricular wall, including the papillary muscles (Figure 14D). In sections of heart tissue hybridised with an

excess of cold probe, no specific labeling over background signal was recorded (Figure 14E).

Apart from the heart, mVRF mRNA signal was present over certain tissues on the outside of the thoracic cage that morphologically resembled brown fat. This was verified with sudan black counterstaining, which showed a strong staining in the same areas (Figure 15A and 15B). In transverse sections of adult mouse spinal cord, the mVRF probes gave a neuronal staining pattern over the gray matter (Figure 15C). Counterstaining with toluidine (Figure 15D) showed that motoneurons in the ventral horn (Figure 15C and 15D), interneurons (Figure 15C) in the deep part of the dorsal horn and around the central canal where to a large extent positive for mVRF mRNA.

### EXAMPLE 7

## EFFECTS OF VEGF AND SOM175 PROTEINS ON CHICK

## SENSORY NEURONS

The effects of VEGF and SOM175 proteins on embryonic day 8 chick sensory neurons were determined using the method of Nurcombe *et al* (1992). The neuronal assay was read at 48 hours using 2000 cells per assay well. The results were obtained using <sup>3</sup>H-thymidine counts. The percentage survival of neurons, neurite outgrowth and average neurite length in  $\mu\text{m}$  were determined using NGF as positive control and various concentrations of VEGF, VEGF in the presence of heparin and VEGF in the presence of heparin and 5  $\mu\text{M}$ , 5'-flurouracil (5FU). 5FU kills glial cells.

25 The results are shown in Figure 16. The results show that VEGF is effective in promoting neuronal survival but that this requires the presence of glial cells. Figure 17 shows the results of the effect of VEGF and SOM175 on three types of chick glia. The glia tested were CNS glia, peripheral glia and CNS oligodendrocytes. Heparin was used as 10  $\mu\text{g/ml}$  in all cultures and the assay was read at 24 hours.

30 Results were measured in  $^3\text{H}$ -thymidine counts using 2000 cells per well.

The results show that for chick central and peripheral neurons, astroglia were markedly stimulated to proliferate by SOM175 in the presence of heparin but that chick oligodendrocytes showed negligible increase in the rate of division.

5

### EXAMPLE 8

#### EFFECTS OF SOM175 PROTEINS ON MOUSE PRIMARY AND CENTRAL NEURONS

The results in Example 7 show that VEGF isoform had an effect on chick primary and central neurons through the agency of the astroglial cells. Similar experiments were repeated in mouse cells.

10

#### Culture conditions

Neuronal and glial cells for all *in vitro* experiments were prepared and cultured according to the techniques described in "Methods in Neurosciences (Vol. 2): Cell Culture" Ed. P.M. Conn, Academic Press, San Diego, 1990, pp33-46 for astroglial cells, pp56-74 for oligodendroglial cells, and pp87-102 for central neurons.

15

Cells were plated onto 24-well culture clusters (Nunc) coated with poly-L-ornithine (0.1 mg/ml, 1h) at a density of 2,000 cells/well. After 48 hours in culture, neurons were counted in the wells under inverted phase light using well established techniques (Maruta *et al.* 1993) and glial cells assessed with [<sup>3</sup>H]thymidine uptake to monitor cell division rates as below. Heparin (10µg/ml, low molecular weight fraction, Sigma Chemical Corp.) was present at all times in the culture media except where noted. The neuronal cultures were supplemented with 5mM 5-fluoro-2-deoxyuridine (Sigma) to suppress background glial growth.

20

25

#### <sup>3</sup>H-Thymidine incorporation assay for glial cell proliferation

The cells were pulsed for 14h with <sup>3</sup>H-thymidine (specific activity 103 µCi/ug) from a stock concentration of 0.1 mCi/ml in standard medium, giving a final incubating volume of 20 µl/well. The contents of the wells were harvested and absorbed onto nitrocellulose paper (Titertek, Flow). Remaining adherent cells were removed by

30

incubation with trypsin/versene (CSL Limited, Victoria, Australia) for 5 min. This procedure was carried out twice. The nitrocellulose discs were washed in a standard Titertek harvester (Flow) using first distilled water, and then methanol. The nitrocellulose discs were dried, scintillation fluid (containing 5% v/v Triton-X) added  
5 and the discs counted on a scintillation counter.

Greatest activity was seen with preparations of SOM175 absent exon 6 (SOM $\Delta$ X6) on mouse astroglial cell cultures, where there was a significant stimulus to their proliferation when delivered in conjunction with heparin (Figure 16). Little stimulus  
10 was given to the proliferation of oligodendroglial cells (Figure 17), and very little discernable potentiation of the survival response of isolated forebrain neurons (Figure 18). The standard deviation on all three graphs for each point was less than 8%.

The viability of neurons can be maintained by promoting glial cell proliferation.  
15 Furthermore, SOM $\Delta$ X6 is a good inducer of astroglial proliferation and may be expressed in conjunction with the formation of astroglial endfeet on central nervous system endothelial cells.

Those skilled in the art will appreciate that the invention described herein is  
20 susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications. The invention also includes all of the steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations of any two or more of said steps or features.

**TABLE 3**  
**Splice junctions of the murine VRF gene**

5	5' UTR* .....	Exon 1 >223bp	CCCAGgtacgtgcgt	Intron I	495bp
	ttccccacagGCCCC	Exon 2 43bp	GAAAGgtaataatag	Intron II	288bp
	ctgcccacagTGGTG	Exon 3 197bp	TGCAGgtaccagggc	Intron III	196bp
	ctgagcacagATCCT	Exon 4 74bp	TGCAGgtgccagccc	Intron IV	182bp
	ctcttttcagACCTA	Exon 5 36bp	GACAGattcttggtg	Intron V	191bp
10	ctcctcctagGGTTG	Exon 6 101bp		(no intron)	
	CCCACTCCAGCCCCA	Exon 7 135bp	TGTAGgtaaggagtc	Intron VI	~2200bp
	cactccccagGTGCC	Exon 8 394bp	AGAGATGGAGACACT		

Uppercase and lowercase letters denote exonic and intronic sequences respectively.

- 15 \* Indicates that the 5' end of exon 1 has not yet been determined.



## BIBLIOGRAPHY

- Adams MD, Soares MB, Kerlavage AR, Fields C, Venter JC, (1993) *Nature Genet.* 4, 373-380.
- Birnstiel ML, Busslinger M and Strub K (1985) *Cell* 41, 349-359.
- Breier G, Albrecht U, Sterrer S and Risau W (1992) *Development* 114, 521-532.
- Chomczynski P and Sacchi N (1987) *Analyt. Biochem.* 162, 156-159.
- Church G and Gilbert W (1984) *Proc. Natl. Acad. Sci. USA* 81, 1991-1995.
- Dagerlind A, Friberg K, Bean AJ and Hokfelt T (1992) *Histochemistry* 98, 39-49.
- Dissen GA, Lara HE, Fabrenbach WH, Costa ME, Ojeda SR, (1994) *Endocrinology* 134, 1146-1154.
- Drinkwater CC, Barker PA, Suter U and Shooter EM (1993) *J. Biol. Chem.*, 268, 23202-23207.
- Drinkwater CC, Suter U, Angst C and Shooter EM (1991) *Proc. Roy. Soc. Lond. (Series B)*, 246, 307-313.
- Ewton DZ & Florini JR (1980) *Endocrinology*, 106: 577-583.
- Ferrara N & Henzel WJ (1989) *Biochem. Biophys. Res. Commun.* 161, 851-858.
- Folkman J & Shing Y (1992) *J. Biol. Chem.* 267, 10931-10934.
- Gospodarowicz D, Abraham JA & Schilling J (1989) *Proc. Natl. Acad. Sci USA* 86, 7311-7315.
- Gospodarowicz D, Weseman J, Morgan JS & Lindstrom J (1976) *J. Cell Biol.*, 70: 395-405.
- Hagg T, Quon D, Higaki J & Varon S (1992) *Neuron*, 8, 145-158.
- Hefti S (1986) *J. Neurosci.* 6, 2155-2162.
- Hendry IA & Campbell J (1976) *J. Neurocytol.*, 5, 351-360.
- Hendry IA, Murphy M, Hilton DJ, Nicola NA & Bartlett PF (1992) *J. Neurosci.* 12, 3427-3434.
- Hunt et al., (1967) *Am. J. Surgery*, 114: 302-307.
- Koch AE, Harlow LA, Haines GK, Amento EP, Unemoti EN, Wong WL, Pope RM, Ferrara N, (1994) *J. Immunol.* 152, 4149-4156.
- Kromer AF (1987) *Science*, 235, 214-216.

- Larsson C, Weber G, Kvanta E, Lewis C, Janson M, Jones C, Glaser T, Evans G, Nordenskjold M, (1992) *Hum. Genet.* **89**, 187-193.
- OZUS Leung DW, Cachianes G, Kuang W-J, Goeddel DV & Ferrara N (1989) *Science* **246**:1306-1309.
- Lowe C, Cornish J, Callon K, Martin TJ & Reid IR (1991) *J. Bone Mineral Res.*, **6**, 1277-1283.
- Lowe C, Cornish J, Martin TJ & Reid IR (1991) *Calcif. Tissue Int.*, **49**, 394-397.
- Martinou JC, Martinou I & Kato AC (1992) *Neuron*, **8**, 737-744.
- Maruta *et al* (1993) *Growth Factors* **8**: 119-134.
- Midy V & Plouet J (1994) *Biochem. Biophys. Res. Commun.*, **199**: 380-386.
- Miles AA & Miles EM (1952) *J. Physiol. (Lond)* **118**:228-257.
- Montesano R, Vassalli JD, Baird A, Guillemin R & Orci, L (1986) *Proc. Natl. Acad. Sci USA*, **83**, 7297-7301.
- Nurcombe *et al* (1992) *Development* **116**: 1175-1183.
- Otto D., Frotscher M & Unsicker K (1989) *J. Neurosci. Res.*, **22**, 83-91.
- Pepper MS, Ferrara N, Orci L, Montesano R. (1991) *Biochem. Biophys. Res. Commun.* **181**, 902-906).
- Rochell JM, Watson ML, Oakey RJ and Seldin MF (1992) *Genomics* **14**, 26-31.
- Roth S & Weston J (1967) *Proc. Natl. Acad. Sci USA*, **58**: 974-980.
- Sambrook J, Fritsch EF, Maniatis T, (1989) *Molecular Cloning: A Laboratory Manual - 2nd Ed.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Santos MA (1991) *Nucleic Acids Res.* **19**, 5442.
- Schilling *et al.*, (1959) *Surgery*, **46**: 702-710.
- Senger DR, Van De Water L, Brown LF, Nagy JA, Yeo KT, Yeo TK, Berse B, Jackman RW, Dvorak AM, Dvorak HF (1993) *Cancer Metastasis Rev.* **12**, 303-324.
- Sharkey AM, Chamock-Jones DS, Boocock CA, Brown KD, Smith SK, (1993) *J. Reprod. Fertil.* **99**, 609-615.
- Sunderkotter C, Steinbrink K, Goebeler M, Bhardway R, Sorg E, (1993) *J. Leukocyt. Biol.* **55**, 410-422.
- Suter U, Angst C, Tien C-L, Drinkwater CC, Lindsay RM and Shooter EM (1992) *J. Neurosci.*, **12**, 306-318.

Tischer E, Mitchell R, Hartman T, Silva M, Gospodarowicz D, Fiddes JC, & Abraham J (1991) *J. Biol. Chem.* 266, 11947-11954.

Williams LR, Varon S, Peterson GM, Wictorin K, Fischer W, Bjorklund A & Gage FH (1986) *Proc. Natl. Acad. Sci. USA* 83, 9231-9235.

Yan Z, Weich HA, Bernart W, Breckwoldt M, Neulen J, (1993) *J. Clin. Endocrinol. Metab.* 77, 1723-1725.

Yip NK, Rich KM, Lampe PA & Johnson EM Jr (1984) *J. Neurosci.*, 4, 2986-2992.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(countries other than US) AMRAD OPERATIONS PTY. LTD.  
(us only) Hayward, N and Weber, G

(ii) TITLE OF INVENTION: A NOVEL GROWTH FACTOR AND A  
GENETIC SEQUENCE ENCODING SAME

(iii) NUMBER OF SEQUENCES: 14

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: DAVIES COLLISON CAVE  
(B) STREET: 1 LITTLE COLLINS STREET  
(C) CITY: MELBOURNE  
(D) STATE: VICTORIA  
(E) COUNTRY: AUSTRALIA  
(F) ZIP: 3000

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: PCT INTERNATIONAL  
(B) FILING DATE: 22-FEB-1996

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: AU PN1457  
(B) FILING DATE: 02-MAR-1995  
(A) APPLICATION NUMBER: AU PN6647  
(B) FILING DATE: 20-NOV-1995  
(A) APPLICATION NUMBER: AU PN7274  
(B) FILING DATE: 22-DEC-1995

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: HUGHES DR, E JOHN L

(C) REFERENCE/DOCKET NUMBER: EJH/EK

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: +61 3 9254 2777

(B) TELEFAX: +61 3 9254 2770

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 649 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 17...589

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TCGGGCCTCC GAAACC ATG AAC TTT CTG CTG TCT TGG GTG CAT TGG AGC	49
Met Asn Phe Leu Leu Ser Trp Val His Trp Ser	
1 5 10	
CTT GCC TTG CTG CTC TAC CTC CAC CAT GCC AAG TGG TCC CAG GCT GCA	97
Leu Ala Leu Leu Leu Tyr Leu His His Ala Lys Trp Ser Gln Ala Ala	
15 20 25	
CCC ATG GCA GAA GGA GGA GGG CAG AAT CAT CAC GAA GTG GTG AAG TTC	145
Pro Met Ala Glu Gly Gly Gly Gln Asn His His Glu Val Val Lys Phe	
30 35 40	
ATG GAT GTC TAT CAG CGC AGC TAC TGC CAT CCA ATC GAG ACC CTG GTG	193
Met Asp Val Tyr Gln Arg Ser Tyr Cys His Pro Ile Glu Thr Leu Val	
45 50 55	
GAC ATC TTC CAG GAG TAC CCT GAT GAG ATC GAG TAC ATC TTC AAG CCA	241
Asp Ile Phe Gln Glu Tyr Pro Asp Glu Ile Glu Tyr Ile Phe Lys Pro	
60 65 70 75	
TCC TGT GTG CCC CTG ATG CGA TGC GGG GGC TGC TGC AAT GAC GAG GGC	289
Ser Cys Val Pro Leu Met Arg Cys Gly Gly Cys Cys Asn Asp Glu Gly	
80 85 90	

CTG GAG TGT GTG CCC ACT GAG GAG TCC AAC ATC ACC ATG CAG ATT ATG Leu Glu Cys Val Pro Thr Glu Glu Ser Asn Ile Thr Met Gln Ile Met 95 100 105	337
CGG ATC AAA CCT CAC CAA GGC CAG CAC ATA GGA GAG ATG AGC TTC CTA Arg Ile Lys Pro His Gln Gly Gln His Ile Gly Glu Met Ser Phe Leu 110 115 120	385
CAG CAC AAC AAA TGT GAA TGC AGA CCA AAG AAA GAT AGA GCA AGA CAA Gln His Asn Lys Cys Glu Cys Arg Pro Lys Lys Asp Arg Ala Arg Gln 125 130 135	433
GAA AAT CCC TGT GGG CCT TGC TCA GAG CGG AGA AAG CAT TTG TTT GTA Glu Asn Pro Cys Gly Pro Cys Ser Glu Arg Arg Lys His Leu Phe Val 140 145 150 155	481
CAA GAT CCG CAG ACG TGT AAA TGT TCC TGC AAA AAC ACA GAC TCG CGT Gln Asp Pro Gln Thr Cys Lys Cys Ser Cys Lys Asn Thr Asp Ser Arg 160 165 170	529
TGC AAG GCG AGG CAG CTT GAG TTA AAC GAA CGT ACT TGC AGA TGT GAC Cys Lys Ala Arg Gln Leu Glu Leu Asn Glu Arg Thr Cys Arg Cys Asp 175 180 185	577
AAG CCG AGG CGG TGAGCCGGGC AGGAGGAAGG AGCCTCCCTC AGCGTTTCGG Lys Pro Arg Arg 190	629
GAACCAGATC TCTCACCAGG	649

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 191 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Asn Phe Leu Leu Ser Trp Val His Trp Ser Leu Ala Leu Leu Leu 1 5 10 15
Tyr Leu His His Ala Lys Trp Ser Gln Ala Ala Pro Met Ala Glu Gly 20 25 30
Gly Gly Gln Asn His His Glu Val Val Lys Phe Met Asp Val Tyr Gln 35 40 45
Arg Ser Tyr Cys His Pro Ile Glu Thr Leu Val Asp Ile Phe Gln Glu 50 55 60

Tyr	Pro	Asp	Glu	Ile	Glu	Tyr	Ile	Phe	Lys	Pro	Ser	Cys	Val	Pro	Leu	
65					70				75						80	
Met	Arg	Cys	Gly	Gly	Cys	Cys	Asn	Asp	Glu	Gly	Leu	Glu	Cys	Val	Pro	
			85						90					95		
Thr	Glu	Glu	Ser	Asn	Ile	Thr	Met	Gln	Ile	Met	Arg	Ile	Lys	Pro	His	
			100					105					110			
Gln	Gly	Gln	His	Ile	Gly	Glu	Met	Ser	Phe	Leu	Gln	His	Asn	Lys	Cys	
		115					120						125			
Glu	Cys	Arg	Pro	Lys	Lys	Asp	Arg	Ala	Arg	Gln	Glu	Asn	Pro	Cys	Gly	
		130				135						140				
Pro	Cys	Ser	Glu	Arg	Arg	Lys	His	Leu	Phe	Val	Gln	Asp	Pro	Gln	Thr	
145					150					155					160	
Cys	Lys	Cys	Ser	Cys	Lys	Asn	Thr	Asp	Ser	Arg	Cys	Lys	Ala	Arg	Gln	
				165				170							175	
Leu	Glu	Leu	Asn	Glu	Arg	Thr	Cys	Arg	Cys	Asp	Lys	Pro	Arg	Arg		
			180					185					190			

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1094 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3..624

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CC ATG AGC CCT CTG CTC CGC CGC CTG CTG CTC GCC GCA CTC CTG CAG	47
Met Ser Pro Leu Leu Arg Arg Leu Leu Leu Ala Ala Leu Leu Gln	
1 5 10 15	
CTG GCC CCC GCC CAG GCC CCT GTC TCC CAG CCT GAT GCC CCT GGC CAC	95
Leu Ala Pro Ala Gln Ala Pro Val Ser Gln Pro Asp Ala Pro Gly His	
20 25 30	
CAG AGG AAA GTG GTG TCA TGG ATA GAT GTG TAT ACT CGC GCT ACC TGC	143
Gln Arg Lys Val Val Ser Trp Ile Asp Val Tyr Thr Arg Ala Thr Cys	
35 40 45	

CAG CCC CGG GAG GTG GTG GTG CCC TTG ACT GTG GAG CTC ATG GGC ACC Gln Pro Arg Glu Val Val Val Pro Leu Thr Val Glu Leu Met Gly Thr 50 55 60	191
GTG GCC AAA CAG CTG GTG CCC AGC TGC GTG ACT GTG CAG CGC TGT GGT Val Ala Lys Gln Leu Val Pro Ser Cys Val Thr Val Gln Arg Cys Gly 65 70 75	239
GGC TGC TGC CCT GAC GAT GGC CTG GAG TGT GTG CCC ACT GGG CAG CAC Gly Cys Cys Pro Asp Asp Gly Leu Glu Cys Val Pro Thr Gly Gln His 80 85 90 95	287
CAA GTC CGG ATG CAG ATC CTC ATG ATC CGG TAC CCG AGC AGT CAG CTG Gln Val Arg Met Gln Ile Leu Met Ile Arg Tyr Pro Ser Ser Gln Leu 100 105 110	335
GGG GAG ATG TCC CTG GAA GAA CAC AGC CAG TGT GAA TGC AGA CCT AAA Gly Glu Met Ser Leu Glu Glu His Ser Gln Cys Glu Cys Arg Pro Lys 115 120 125	383
AAA AAG GAC AGT GCT GTG AAG CCA GAC AGG GCT GCC ACT CCC CAC CAC Lys Lys Asp Ser Ala Val Lys Pro Asp Arg Ala Ala Thr Pro His His 130 135 140	431
CGT CCC CAG CCC CGT TCT GTT CCG GGC TGG GAC TCT GCC CCC GGA GCA Arg Pro Gln Pro Arg Ser Val Pro Gly Trp Asp Ser Ala Pro Gly Ala 145 150 155	479
CCC TCC CCA GCT GAC ATC ACC CAT CCC ACT CCA GCC CCA GGC CCC TCT Pro Ser Pro Ala Asp Ile Thr His Pro Thr Pro Ala Pro Gly Pro Ser 160 165 170 175	527
GCC CAC GCT GCA CCC AGC ACC ACC AGC GCC CTG ACC CCC GGA CCT GCC Ala His Ala Ala Pro Ser Thr Thr Ser Ala Leu Thr Pro Gly Pro Ala 180 185 190	575
GCT GCC GCT GCC GAC GCC GCA GCT TCC TCC GTT GCC AAG GGC GGG GCT T Ala Ala Ala Ala Asp Ala Ala Ala Ser Ser Val Ala Lys Gly Gly Ala 195 200 205	624
AGAGCTCAAC CCAGACACCT GCAGGTGCCG GAAGCTGCGA AGGTGACACA TGGCTTTTCA	684
GACTCAGCAG GGTGACTTGC CTCAGAGGCT ATATCCCAGT GGGGGAACAA AGGGGAGCCT	744
GGTAAAAAAC AGCCAAGCCC CCAAGACCTC AGCCCAGGCA GAAGCTGCTC TAGGACCTGG	804
GCCTCTCAGA GGGCTCTTCT GCCATCCCTT GTCTCCCTGA GGCCATCATC AAACAGGACA	864
GAGTTGGAAG AGGAGACTGG GAGGCAGCAA GAGGGGTCAC ATACCAGCTC AGGGGAGAAT	924
GGAGTACTGT CTCAGTTTCT AACCACCTCTG TGCAAGTAAG CATCTTACAA CTGGCTCTTC	984
CTCCCCTCAC TAAGAAGACC CAAACCTCTG CATAATGGGA TTTGGGCTTT GGTACAAGAA	1044
CTGTGACCCC CAACCCTGAT AAAAGAGATG GAAGGAAAAA AAAAAAAAAA	1094



(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 207 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Ser Pro Leu Leu Arg Arg Leu Leu Leu Ala Ala Leu Leu Gln Leu  
1 5 10 15  
Ala Pro Ala Gln Ala Pro Val Ser Gln Pro Asp Ala Pro Gly His Gln  
20 25 30  
Arg Lys Val Val Ser Trp Ile Asp Val Tyr Thr Arg Ala Thr Cys Gln  
35 40 45  
Pro Arg Glu Val Val Val Pro Leu Thr Val Glu Leu Met Gly Thr Val  
50 55 60  
Ala Lys Gln Leu Val Pro Ser Cys Val Thr Val Gln Arg Cys Gly Gly  
65 70 75 80  
Cys Cys Pro Asp Asp Gly Leu Glu Cys Val Pro Thr Gly Gln His Gln  
85 90 95  
Val Arg Met Gln Ile Leu Met Ile Arg Tyr Pro Ser Ser Gln Leu Gly  
100 105 110  
Glu Met Ser Leu Glu Glu His Ser Gln Cys Glu Cys Arg Pro Lys Lys  
115 120 125  
Lys Asp Ser Ala Val Lys Pro Asp Arg Ala Ala Thr Pro His His Arg  
130 135 140  
Pro Gln Pro Arg Ser Val Pro Gly Trp Asp Ser Ala Pro Gly Ala Pro  
145 150 155 160  
Ser Pro Ala Asp Ile Thr His Pro Thr Pro Ala Pro Gly Pro Ser Ala  
165 170 175  
His Ala Ala Pro Ser Thr Thr Ser Ala Leu Thr Pro Gly Pro Ala Ala  
180 185 190  
Ala Ala Ala Asp Ala Ala Ala Ser Ser Val Ala Lys Gly Gly Ala  
195 200 205

- 42 -

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 993 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 3..566

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CC ATG AGC CCT CTG CTC CGC CGC CTG CTG CTC GCC GCA CTC CTG CAG	47
Met Ser Pro Leu Leu Arg Arg Leu Leu Leu Ala Ala Leu Leu Gln	
1 5 10 15	
CTG GCC CCC GCC CAG GCC CCT GTC TCC CAG CCT GAT GCC CCT GGC CAC	95
Leu Ala Pro Ala Gln Ala Pro Val Ser Gln Pro Asp Ala Pro Gly His	
20 25 30	
CAG AGG AAA GTG GTG TCA TGG ATA GAT GTG TAT ACT CGC GCT ACC TGC	143
Gln Arg Lys Val Val Ser Trp Ile Asp Val Tyr Thr Arg Ala Thr Cys	
35 40 45	
CAG CCC CGG GAG GTG GTG GTG CCC TTG ACT GTG GAG CTC ATG GGC ACC	191
Gln Pro Arg Glu Val Val Val Pro Leu Thr Val Glu Leu Met Gly Thr	
50 55 60	
GTG GCC AAA CAG CTG GTG CCC AGC TGC GTG ACT GTG CAG CGC TGT GGT	239
Val Ala Lys Gln Leu Val Pro Ser Cys Val Thr Val Gln Arg Cys Gly	
65 70 75	
GGC TGC TGC CCT GAC GAT GGC CTG GAG TGT GTG CCC ACT GGG CAG CAC	287
Gly Cys Cys Pro Asp Asp Gly Leu Glu Cys Val Pro Thr Gly Gln His	
80 85 90 95	
CAA GTC CGG ATG CAG ATC CTC ATG ATC CGG TAC CCG AGC AGT CAG CTG	335
Gln Val Arg Met Gln Ile Leu Met Ile Arg Tyr Pro Ser Ser Gln Leu	
100 105 110	
GGG GAG ATG TCC CTG GAA GAA CAC AGC CAG TGT GAA TGC AGA CCT AAA	383
Gly Glu Met Ser Leu Glu Glu His Ser Gln Cys Glu Cys Arg Pro Lys	
115 120 125	
AAA AAG GAC AGT GCT GTG AAG CCA GAT AGC CCC AGG CCC CTC TGC CCA	431
Lys Lys Asp Ser Ala Val Lys Pro Asp Ser Pro Arg Pro Leu Cys Pro	
130 135 140	

- 43 -

CGC TGC ACC CAG CAC CAC CAG CGC CCT GAC CCC CGG ACC TGC CGC TGC 479  
 Arg Cys Thr Gln His His Gln Arg Pro Asp Pro Arg Thr Cys Arg Cys  
 145 150 155

CGC TGC CGA CGC CGC AGC TTC CTC CGT TCC CAA GGG CGG GGC TTA GAC 527  
 Arg Cys Arg Arg Ser Phe Leu Arg Cys Gln Gly Arg Gly Leu Glu  
 160 165 170 175

CTC AAC CCA GAC ACC TGC AGG TGC CGG AAG CTG CGA AGG TGACACATGG 576  
 Leu Asn Pro Asp Thr Cys Arg Cys Arg Lys Leu Arg Arg  
 180 185

CTTTTCAGAC TCAGCAGGGT GACTTGCCCTC AGAGGCTATA TCCCAGTGGG GGAACAAAGG 636

GGAGCCTGGT AAAAAACAGC CAAGCCCCCA AGACCTCAGC CCAGGCAGAA GCTGCTCTAG 696

GACCTGGGCC TCTCAGAGGG CTCTTCTGCC ATCCCTTGTC TCCCTGAGGC CATCATCAAA 756

CAOGACAGAG TTGGAAGAGG AGACTGGGAG GCACCAAGAG GGGTCACATA CCAGCTCAGG 816

GGAGAATGGA GTACTGTCTC AGTTTCTAAC CACTCTGTGC AAGTAAGCAT CTTACAACTC 876

GCTCTTCCTC CCCTCACTAA GAAGACCCAA ACCTCTGCAT AATGGGATTT GGCCTTTTCT 936

ACAAGAACTG TGACCCCCAA CCCTGATAAA AGAGATGGAA GGAAAAAAA AAAAAAA 997

## (2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 188 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Ser Pro Leu Leu Arg Arg Leu Leu Leu Ala Ala Leu Leu Gln Leu  
 1 5 10 15

Ala Pro Ala Gln Ala Pro Val Ser Gln Pro Asp Ala Pro Gly His Gln  
 20 25 30

Arg Lys Val Val Ser Trp Ile Asp Val Tyr Thr Arg Ala Thr Cys Gln  
 35 40 45

Pro Arg Glu Val Val Val Pro Leu Thr Val Glu Leu Met Gly Thr Val  
 50 55 60

Ala Lys Gln Leu Val Pro Ser Cys Val Thr Val Gln Arg Cys Gly Gly  
 65 70 75 80

Cys Cys Pro Asp Asp Gly Leu Glu Cys Val Pro Thr Gly Gln His Gln  
85 90 95  
Val Arg Met Gln Ile Leu Met Ile Arg Tyr Pro Ser Ser Gln Leu Gly  
100 105 110  
Glu Met Ser Leu Glu Glu His Ser Gln Cys Glu Cys Arg Pro Lys Lys  
115 120 125  
Lys Asp Ser Ala Val Lys Pro Asp Ser Pro Arg Pro Leu Cys Pro Arg  
130 135 140  
Cys Thr Gln His His Gln Arg Pro Asp Pro Arg Thr Cys Arg Cys Arg  
145 150 155 160  
Cys Arg Arg Arg Ser Phe Leu Arg Cys Gln Gly Arg Gly Leu Glu Leu  
165 170 175  
Asn Pro Asp Thr Cys Arg Cys Arg Lys Leu Arg Arg  
180 185

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 858 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3..431

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CC ATG AGC CCT CTG CTC CGC CGC CTG CTG CTC GCC GCA CTC CTG CAG	47
Met Ser Pro Leu Leu Arg Arg Leu Leu Leu Ala Ala Leu Leu Gln	
1 5 10 15	
CTG GCC CCC GCC CAG GCC CCT GTC TCC CAG CCT GAT GCC CCT GGC CAC	95
Leu Ala Pro Ala Gln Ala Pro Val Ser Gln Pro Asp Ala Pro Gly His	
20 25 30	
CAG AGG AAA GTG GTG TCA TGG ATA GAT GTG TAT ACT CGC GCT ACC TGC	143
Gln Arg Lys Val Val Ser Trp Ile Asp Val Tyr Thr Arg Ala Thr Cys	
35 40 45	
CAG CCC CGG GAG GTG GTG GTG CCC TTG ACT GTG GAG CTC ATG GGC ACC	191
Gln Pro Arg Glu Val Val Val Pro Leu Thr Val Glu Leu Met Gly Thr	
50 55 60	

- 45 -

GTG GCC AAA CAG CTG GTG CCC AGC TGC GTG ACT GTG CAG CCG TGT GGT Val Ala Lys Gln Leu Val Pro Ser Cys Val Thr Val Gln Arg Cys Gly 65 70 75	239
GGC TCC TGC CCT GAC GAT GGC CTG GAG TCT GTG CCC ACT GGG CAG CAC Gly Cys Cys Pro Asp Asp Gly Leu Glu Cys Val Pro Thr Gly Gln His 80 85 90 95	207
CAA GTC CGG ATG CAG ATC CTC ATG ATC CCG TAC CCG AGC AGT CAG CTG Gln Val Arg Met Gln Ile Leu Met Ile Arg Tyr Pro Ser Ser Gln Leu 100 105 110	115
GGG GAG ATG TCC CTG GAA GAA CAC AGC CAG TGT GAA TGC AGA CCT AAA Gly Glu Met Ser Leu Glu Glu His Ser Gln Cys Glu Cys Arg Pro Lys 115 120 125	383
AAA AAG CAC AGT CCT GTG AAG CCA GAT AGG TGC CCG AAG CTG CGA AGC Lys Lys Asp Ser Ala Val Lys Pro Asp Arg Cys Arg Lys Leu Arg Arg 130 135 140	431
TGACACATGG CTTTTCAGAC TCAGCAGGGT GACTTGCCCTC AGAGGCTATA TCCCAGTGGG GGAACAAAGG GGAGCCTGGT AAAAAACAGC CAAGCCCCCA AGACCTCAGC CCAGGCACAA OCTGCTCTAG GACCTGGGCC TCTCAGAGGG CTCTTCTGCC ATCCCTTGTC TCCTGAGGC CATCATCAAA CAGGACAGAG TTGGAAGAGG AGACTGGGAG GCAGCAAGAG GGCTCACATA CCAGCTCAGG GGAGAATGGA GTACTGTCTC AGTTTCTAAC CACTCTGTGC AAGTAAGCAT CTTACAACTG GCTCTTCCTC CCCTCACTAA GAAGACCCAA ACCTCTGCAT AATGGGATTT GGGCTTTGGT ACAAGAACTG TGACCCCCAA CCCTGATAAA AGAGATGGAA GGAAAAAAAAA AAAAAAA	491 551 611 671 731 791 851 850

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 143 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
Met Ser Pro Leu Leu Arg Arg Leu Leu Leu Ala Ala Leu Leu Gln Leu
 1             5             10             15
Ala Pro Ala Gln Ala Pro Val Ser Gln Pro Asp Ala Pro Gly His Gln
          20             25             30
Arg Lys Val Val Ser Trp Ile Asp Val Tyr Thr Arg Ala Thr Cys Gln
          35             40             45
Pro Arg Glu Val Val Val Pro Leu Thr Val Glu Leu Met Gly Thr Val
          50             55             60
Ala Lys Gln Leu Val Pro Ser Cys Val Thr Val Gln Arg Cys Gly Gly
 65             70             75             80
Cys Cys Pro Asp Asp Gly Leu Glu Cys Val Pro Thr Gly Gln His Gln
          85             90             95
Val Arg Met Gln Ile Leu Met Ile Arg Tyr Pro Ser Ser Gln Leu Gly
          100            105            110
Glu Met Ser Leu Glu Glu His Ser Gln Cys Glu Cys Arg Pro Lys Lys
          115            120            125
Lys Asp Ser Ala Val Lys Pro Asp Arg Cys Arg Lys Leu Arg Arg
          130            135            140
```

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 910 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3..305

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CC ATG AGC CCT CTG CTC CGC CGC CTG CTG CTC GCC GCA CTC CTG CAG	47
Met Ser Pro Leu Leu Arg Arg Leu Leu Leu Ala Ala Leu Leu Gln	
1 5 10 15	
CTG GCC CCC GCC CAG GCC CCT GTC TCC CAG CCT GAT GCC CCT GGC CAC	95
Leu Ala Pro Ala Gln Ala Pro Val Ser Gln Pro Asp Ala Pro Gly His	
20 25 30	
CAG AGG AAA GTG GTG TCA TGG ATA GAT GTG TAT ACT CGC GCT ACC TGC	143
Gln Arg Lys Val Val Ser Trp Ile Asp Val Tyr Thr Arg Ala Thr Cys	
35 40 45	
CAG CCC CGG GAG GTG GTG GTG CCC TTG ACT GTG GAG CTC ATG GGC ACC	191
Gln Pro Arg Glu Val Val Val Pro Leu Thr Val Glu Leu Met Gly Thr	
50 55 60	
GTG GCC AAA CAG CTG GTG CCC AGC TGC GTG ACT GTG CAG CGC TGT GGT	239
Val Ala Lys Gln Leu Val Pro Ser Cys Val Thr Val Gln Arg Cys Gly	
65 70 75	
GGC TGC TGC CCT GAC GAT GGC CTG GAG TGT GTG CCC ACT GGG CAG CAC	287
Gly Cys Cys Pro Asp Asp Gly Leu Glu Cys Val Pro Thr Gly Gln His	
80 85 90 95	
CAA GTC CGG ATG CAG ACC TAAAAAAG GACAGTGCTG TGAAGCCAGA	335
Gln Val Arg Met Gln Thr	
100	
CAGGGCTGCC ACTCCCCACC ACCGTCCCCA GCCCCGTTCT GTTCCGGGCT GGGACTCTGC	395
CCCCGGAGCA CCTCCCCAG CTGACATCAC CCATCCCACT CCAGCCCCAG GCCCCCTCTGC	455
CCACGCTGCA CCCAGCACCA CCAGCGCCCT GACCCCCGGA CTTGCCGCTG CCGCTGCCGA	515
CGCCGCAGCT TCCTCCGTTG CCAAGGGCGG GGCTTAGAGC TCAACCCAGA CACCTGCAGG	575
TGCCGGAAGC TGCGAAGGTG ACACATGGCT TTTCAGACTC AGCAGGGTGA CTTGCCTCAG	635
AGGCTATATC CCAAGTGGGA ACAAAGAGGA GCCTGGTAAA AAACAGCCAA GCCCCCAAGA	695

- 48 -

CCTCAGCCCA GGCAGAAGCT GCTCTAGGAC CTGGGCCTCT CAGAGGGCTC TTCTGCCATC 755  
 CCTTGTCTCC CTGAGGCCAT CATCAAACAG GACAGAGTTG GAAGAGGAGA CTGGGAGGCA 815  
 GCAAGAGGGG TCACATACCA GCTCAGGGGA GAATGGAGTA CTGTCTCAGT TTCTAACCAC 875  
 TCTGTGCAAG TAAGCATCTT ACAACTGGCT CTTCC 910

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 101 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Ser Pro Leu Leu Arg Arg Leu Leu Leu Ala Ala Leu Leu Gln Leu  
 1 5 10 15  
 Ala Pro Ala Gln Ala Pro Val Ser Gln Pro Asp Ala Pro Gly His Gln  
 20 25 30  
 Arg Lys Val Val Ser Trp Ile Asp Val Tyr Thr Arg Ala Thr Cys Gln  
 35 40 45  
 Pro Arg Glu Val Val Val Pro Leu Thr Val Glu Leu Met Gly Thr Val  
 50 55 60  
 Ala Lys Gln Leu Val Pro Ser Cys Val Thr Val Gln Arg Cys Gly Gly  
 65 70 75 80  
 Cys Cys Pro Asp Asp Gly Leu Glu Cys Val Pro Thr Gly Gln His Gln  
 85 90 95  
 Val Arg Met Gln Thr  
 100



(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 42 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ACCACCACCT CCCTGGGCTG GCATGTGGCA CGTGCATAAA CG

42

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 42 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

AGTTGTTTGA CCACATTGCC CATGAGTTCC ATGCTCAGAG GC

42

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 38 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GATCCTGGGG CTGGAGTGGG ATGGATGATG TCAGCTGG

38

- 50 -

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GCGGGCAGAG GATCCTGGGG CTGTCTGGCC TCACAGCACT

40